



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

DOES AGGREGATION LEAD TO BIASED INFERENCES?  
AN EMPIRICAL ANALYSIS OF THE ADOPTION OF OIL-  
PUMPKIN CULTIVATION IN AUSTRIA AT THE FARM- AND  
MUNICIPALITY-LEVEL

Andreas Niedermayr and Jochen Kantelhardt

a.niedermayr@boku.ac.at

Institut für Agrar- und Forstökonomie, Universität für Bodenkultur Wien,  
Feistmantelstraße 4, 1180 Wien, Österreich



2017

***Posterpräsentation anlässlich der 57. Jahrestagung der GEWISOLA***  
*(Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V.)*  
***und der 27. Jahrestagung der ÖGA***  
*(Österreichische Gesellschaft für Agrarökonomie)*  
***„Agrar- und Ernährungswirtschaft zwischen Ressourceneffizienz und***  
***gesellschaftlichen Erwartungen“***  
***Weihenstephan, 13. bis 15. September 2017***

Copyright 2017 by authors. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

# **DOES AGGREGATION LEAD TO BIASED INFERENCES? AN EMPIRICAL ANALYSIS OF THE ADOPTION OF OIL-PUMPKIN CULTIVATION IN AUSTRIA AT THE FARM- AND MUNICIPALITY-LEVEL**

*Andreas Niedermayr and Jochen Kantelhardt*

## **Summary**

The aim of this study is to assess, whether estimation of the same innovation-adoption model at farm- and municipality-level results in an ecological fallacy, meaning that based on aggregated data, one would make inverse inferences about the driving forces influencing the adoption decision at the farm level. The adoption of an emerging alternative crop in Austria, the Styrian Oil Pumpkin, serves as an applied example. Our findings indicate the presence of an ecological fallacy. We therefore propose further research, which could consist of Monte Carlo simulations in order to analyse sensitivity of results with respect to the degree of aggregation.

## **Keywords**

Innovation adoption, ecological fallacy, SLX-Tobit model, Styrian Oil Pumpkin, PGI.

## **1 Introduction**

Empirical innovation adoption studies are interested in estimating the effect of various driving forces on the adoption of innovations. As innovation adoption often occurs in spatial clusters, the notion of spatial spillover effects (adoption of an innovation in one place affects adoption in nearby places) is of particular interest in this context. In order to analyse potential spatial spillover effects, spatially explicit data of the whole population of interest (e.g. farms) is needed. As complete census data at the farm level is hardly available and limited resources prevent large-scale surveys of the whole farm population, researchers mostly use aggregated data (e.g. Schmidtner et al., 2012; Garrett et al., 2013; Niedermayr et al., 2016). However, this approach may result in an “ecological fallacy” (Openshaw, 1984; Anselin, 2002), meaning that the usage of aggregated data to make inferences about a process that happens at the farm level (the adoption decision) may lead to inverse inferences about the true relationship of interest. While limited research that compares the outcomes of such studies at different aggregation levels exists (e.g. Schmidtner et al., 2015), we are not aware of any empirical analysis comparing aggregated- and farm-level results. The aim of this study is therefore to assess, whether aggregation could lead to such an ecological fallacy. The adoption of oil-pumpkin cultivation in an Austrian case study region serves as an applied example.

## **2 Data and Methods**

For the regression analysis, we use previously unavailable, spatially explicit cross-sectional data from 2010 of roughly 7,726 farms in a case study region in Lower Austria (BMLFUW, 2016), where the implementation of a protected geographical indication for Styrian Pumpkin Seed Oil triggered a dynamic development of oil-pumpkin cultivation (Niedermayr et al., 2016). Because of censoring in our dependent variable (share of arable land, cultivated with oil pumpkin), we estimate a Tobit model and extend it to a Spatial Lag of X (SLX) Tobit model. In a SLX model, spatial lags of the independent variables, reflecting for each observation the average value of the respective independent variable of neighbouring observations, are added as further independent variables. This allows estimating potential spatial spillover effects of the independent variables on adoption (Halleck Vega and Elhorst, 2015), reflecting e.g. shared resources for oil-pumpkin cultivation. The independent variables in our model describe natural conditions, availability of oil-pumpkin specific infrastructure, production- marketing- and

policy- related factors, social, temporal and spatial factors. We directly aggregate the farm-level data to the municipality-level, in order to rule out any other sources of influence on the results.

### 3 Preliminary results

Table 1 shows the partial effect at the average (PEA) of the independent variables. While most signs of the significant variables do not change, when comparing municipality- and farm-level results, there are also differences. We briefly illustrate the issue with the variable direct marketing, while noting that a similar line of argument is also possible for others (e.g. the spatial lag variables). Although, direct marketing is beneficial for oil-pumpkin cultivation from a theoretical point of view, the model based on municipality-level data shows a negative relationship. Most likely, at the municipality level, the presence of direct marketing farms, which do not cultivate oil pumpkin, leads to a bias of the true relationship of interest. Such potential ecological fallacies could also be present in comparable studies and are in our case overcome by an analysis at the farm-level. However, the scarce availability of spatially explicit farm-level data is not likely to change in the near future, ruling out this option as a general solution. We therefore propose further research, which could include Monte Carlo simulations in order to analyse the sensitivity of results with respect to the degree of aggregation.

**Table 1: Comparison of marginal effects at the municipality- and farm level**

<b>Independent Variables</b>	<b>Municipality level</b>		<b>Farm level</b>	
Soil-quality index	<b>0,04</b>	<b>n.s.</b>	<b>0,01</b>	<b>***</b>
Distance to nearest drying facility for pumpkin seeds	-0,20	<b>***</b>	-0,10	<b>***</b>
Livestock density	-0,11	<b>*</b>	-0,01	<b>**</b>
Log(farm size)	-0,010	<b>*</b>	-0,003	<b>**</b>
Log(UBAG subsidy for arable land)	-0,000	<b>n.s.</b>	0,000	<b>n.s.</b>
Log(arable land)	<b>0,005</b>	<b>n.s.</b>	<b>0,006</b>	<b>***</b>
Temporal lag of oil-pumpkin share	1,00	<b>***</b>	0,09	<b>***</b>
Direct marketing	<b>-0,06</b>	<b>*</b>	<b>0,15</b>	<b>*</b>
Organic farming	0,09	<b>***</b>	0,95	<b>***</b>
Agricultural education	0,03	<b>**</b>	0,08	<b>*</b>
WX of Direct marketing	-0,05	<b>n.s.</b>	-0,39	<b>n.s.</b>
WX of Organic farming	<b>0,003</b>	<b>n.s.</b>	<b>0,45</b>	<b>**</b>
WX of Agricultural education	<b>-0,01</b>	<b>n.s.</b>	<b>0,34</b>	<b>*</b>

Source: own calculations, data from BMLFUW (2016). Note: the PEAs of the three log-transformed independent variables have been divided by 100 so that a change of x by 1% can be interpreted as a percentage-point change of y; the 3 variables “direct marketing”, “organic farming” and “agricultural education” are shares of all farms/farmers at the municipality level and dummy variables at the farm level; spatial-lag variables are denoted by the prefix “WX of”; **\*\*\***, **\*\*** and **\*** and denote significance at the 1%, 5% and 10% levels. **n.s.**=not significant

### References

**Anselin, L.**, 2002. Under the hood Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27 (3), 247–267.; **BMLFUW**, 2016. Integrated Administration and Control System Database.; **Garrett, R.D., Lambin, E.F., Naylor, R.L.**, 2013. Land institutions and supply chain configurations as determinants of soybean planted area and yields in Brazil. *Land Use Policy* 31, 385–396.; **Halleck Vega, S., Elhorst, J.P.**, 2015. The SLX Model. *Journal of Regional Science* 55 (3), 339–363.; **Niedermayr, A., Kapfer, M., Kantelhardt, J.**, 2016. Regional heterogeneity and spatial interdependence as determinants of the cultivation of an emerging alternative crop: The case of the Styrian Oil Pumpkin. *Land Use Policy* 58, 276–288.; **Openshaw, S.**, 1984. Ecological Fallacies and the Analysis of Areal Census Data. *Environment and Planning A* 16 (1), 17–31.; **Schmidtner, E., Lippert, C., Dabbert, S.**, 2015. Does Spatial Dependence Depend on Spatial Resolution?: An Empirical Analysis of Organic Farming in Southern Germany. *German Journal of Agricultural Economics* 64 (3), 175–191.; **Schmidtner, E., Lippert, C., Engler, B., Haring, A.M., Aurbacher, J., Dabbert, S.**, 2012. Spatial distribution of organic farming in Germany: Does neighbourhood matter? *Eur Rev Agric Econ* 39 (4), 661–683.